Available online at www.sciencedirect.com
ScienceDirect
Journal of Forensic and Legal Medicine 15 (2008) 519–523
JOURNAL OF FORENSIC AND LEGAL MEDICINE
www.elsevier.com/jflm

Short Report

Population data on 10 non-CODIS STR loci in Japanese population using a newly developed multiplex PCR system

H. Asamura PhD MD (Senior Assistant Professor) *,
M. Ota PhD (Senior Assistant Professor), H. Fukushima PhD MD (Professor)

Department of Legal Medicine, Shinshu University School of Medicine, Asahi 3-1-1, Matsumoto, Nagano 390-8621, Japan

Received 21 February 2008; accepted 9 April 2008
Available online 12 June 2008

Abstract

This paper describes a newly devised autosomal short tandem repeat (STR) multiplex polymerase chain reaction (PCR) systems for 10 loci (D1S1656, D2S1353, D8S1132, D12S1090, D14S608, D18S535, D19S253, D20S480, D21S226, and D22S689) unlinked to the core STR loci (non-CODIS loci). Of 252 samples taken from the Japanese population, PCR products ranged in length from 107 bp to 319 bp. No significant deviations from Hardy–Weinberg equilibrium were observed at any of the 10 loci. The accumulated power of discrimination and power of exclusion for the 10 loci were 0.999999999998 and 0.99991, respectively. We conclude that the present multiplex system for the 10 non-CODIS loci represents a powerful tool for forensic applications.
© 2008 Elsevier Ltd and FFLM. All rights reserved.

Keywords: STR; Non-CODIS loci; Multiplex; Japanese population data

Introduction

Analyses of short tandem repeat (STR) markers are an essential component of forensic genetics. Since the 13 core combined DNA index system (CODIS) STR loci were used for a DNA database of felons, several commercially available multiplex polymerase chain reaction (PCR) kits have been developed for analysis of core STR loci and have now entered wide use. On the other hand, many highly polymorphic STR loci exist, which are unlinked to the core STR loci (non-CODIS loci). Analysis of such polymorphic non-CODIS loci in addition to the core CODIS loci may increase the power of exclusion in complex kinship testing involving deficient paternity cases. The present study sought to develop a new multiplex PCR system for simultaneous typing of highly polymorphic non-CODIS loci. We selected 10 polymorphic non-CODIS loci (D1S1656, D2S1353, D8S1132, D12S1090, D14S608, D18S535, D19S253, D20S480, D21S226, and D22S689) from the literature, and surveyed their efficacy for forensic practice.

Materials and methods

Informed consent was obtained and blood samples collected from 252 healthy unrelated individuals residing in Japan, and additionally 24 daughters and 21 sons (45 family trios). Parents of the children were included among the 252 subjects, and the paternal and maternal relationships of the families were confirmed by autosomal STRs genotyping using the AmpflSTR Identifiler Kit (Applied Biosystems, Foster City, CA). DNA was extracted from whole blood using the nucleic acid isolation system Quick Gene-800 (FUJIFILM, Tokyo, Japan) according to the protocol described.

Multiplex PCR for amplification of the 10 non-CODIS STR loci was performed using the fluorescent dye labeled primer sets listed in Table 1. PCR reactions were performed in a total volume of 15 μl containing of 2.0 ng of genomic DNA, 1x GeneAmp PCR buffer, 1.5 mM of $MgCl_2$,

* Corresponding author. Tel.: +81 0263 37 3218; fax: +81 0263 37 3084.
E-mail address: asamura@sch.md.shinshu-u.ac.jp (H. Asamura).

1752-928X/$ - see front matter © 2008 Elsevier Ltd and FFLM. All rights reserved.
doi:10.1016/j.jflm.2008.04.001

Table 1
Primer sequences used in this study

STR locus	Primer sequences		Dye label	Reference
	Forward primer	Reverse primer[a]		
D1S1656	5′-GCTCAAGGGTCAACTGTAT-3′	5′-AAATAGAATCACTAGGGAA-3′	VIC	Newly designed[b]
D2S1353	5′-AGCCAGGGACATTGCTTAA-3′	5′-CCATTAGAGCAGGATTTGT-3′	PET	Newly designed[b]
D8S1132	5′-GGCTAGGAAAGGTTAGTGGC-3′	5′-CCCTCTCTCTTTCGAGCAAT-3′	NED	GDB[c]
D12S1090	5′-AAGCCCAAAGATGTAAGGCT-3′	5′-ACCAACCTAGGAAACACAGT-3′	VIC	[7]
D14S608	5′-TAAAGGTTTATCCATGCTGTAGC-3′	5′-ACGTGGTACAGGTAGATAAATGG-3′	PET	GDB
D18S535	5′-TCATGTGACAAAAGCCACAC-3′	5′-AGACAGAAATATAGATGAGAATGCA-3′	FAM	GDB
D19S253	5′-ATAGACAGACAGACGGACTG-3′	5′-GGGAGTGGAGATTACCCCT-3′	FAM	GDB
D20S480	5′-GTGGTGAACACCAATAAATGG-3′	5′-AAGCAAATAAAACCAATAAACTCG-3′	NED	GDB
D21S226	5′-GGAAACCACTCTAAGACATA-3′	5′-ATTTTTTTTCCTTTGGCTGACA-3′	PET	Newly designed[d]
D22S689	5′-TATGTACAGACCTGCAACTTGC-3′	5′-CCTGCCTGCCTATCTATCTG-3′	NED	GDB

[a] A particular sequence of seven bases, a tailed primer (Applied Biosystems), was added to the 5′ end of all reverse primers to promote adenylation.
[b] Newly designed primer set in this study.
[c] Genome Database (http://www.gdb.org).
[d] Newly designed reverse primer in this study. Forward primer is shown in the GDB.

200 μM of each deoxyribonucleotide triphosphate (dNTP, GeneAmp dNTP MIX), each primer set, and 1.25 U of AmpliTaq Gold DNA polymerase (Applied Biosystems). The multiplex PCR mixture contained 0.25 μM primer sets of D18S535 (fluorescent dye labeled with 6FAM), 0.12 μM D19S253 (6FAM), 0.5 μM D1S1656 (VIC), 0.25 μM D12S1090 (VIC), 0.25 μM D8S1132 (NED), 0.65 μM D22S689 (NED), 1.0 μM D20S480 (NED), 0.95 μM D21S226 (PET), 1.0 μM D2S1353 (PET), and 0.4 μM D14S608 (PET). The cycling programs consisted of pre-denaturation at 95 °C for 11 min, followed by 30 cycles of denaturing at 94 °C for 30 s, annealing at 56 °C for 30 s, extension at 72 °C for 45 s, and a final extension at 70 °C for 60 min with a GeneAmp 9700 (Applied Biosystems) in 9600 mode. Electrophoresis was performed with an ABI 3100-Avant Genetic Analyzer (Applied Biosystems) in which 1.0 μl of PCR product was mixed with 15 μl Hi-Di formamide and 0.2 μl of the GeneScan-500LIZ size standard. GeneMapper ID v3.2 Software was used to analyze the data. Allelic typing was based on sequenced allelic ladders. All fragments contained in allelic ladder were sequenced to calibrate the number of repeats. Non-labeled primer sets, which had the same sequences as given in Table 1, were used for sequencing. The target alleles, re-amplified after extraction using the crush and soak method,[1] were used as templates. Sequencing reactions were performed with the same primers using the BigDye v1.1 Ready Reaction Kit (Applied Biosystems). Following electrophoresis with the ABI 3100-Avant Genetic Analyzer (Applied Biosystems), sequencing was determined with sequencing analysis software 5.1.1 (Applied Biosystems). Genotypes for 252 Japanese individuals and commercially available cell lines, including 9947A (Applied Biosystems), K562 and 9948 DNA (Promega, Madison, WI) were typed using the allelic ladder. To determine the minimum quantity of DNA required to achieve reliable results [peak highs at all loci exceeded 100 relative fluorescence units (RFU)] with the multiplex systems, we used aliquots of the cell line 9947A DNA (50, 100, 200, 300, 400, 500, 750 pg, 1, 5, and 10 ng).

The allele frequency at each locus was calculated from the numbers of alleles in the 252 Japanese individuals. Various forensic parameters were calculated using the PowerStats v1.2 software package (Promega). Hardy–Weinberg equilibrium and population differentiation were determined by an exact test using the GENEPOP (version 3.4) software package designed by Raymond and Rousset.[2]

Table 2
Information on the STR loci in this study and a result of genotype of cell line DNA

STR locus	Product length (bp)[a]	Allele range	Number of alleles	Genotype of cell line DNA			Chromosomal location
				K562	9947A	9948	
D1S1656	129–164	11–19.3	13	15, 16	18.3, 18.3	14, 17	1q42
D2S1353	150–180	8–18	11	10, 12	12, 14	12, 14	2q24
D8S1132	142–178	16–25	10	20, 24	19, 21	20, 24	8q23
D12S1090	213–297	8–29	20	21, 25	19, 20	21, 22	12q12
D14S608	201–233	6–14	9	7, 7	7, 11	12, 12	14q12
D18S535	138–174	9–18	10	13, 14	13, 15	13, 15	18q12
D19S253	213–249	6–15	10	11, 12	7, 8	11, 12	19p13.1
D20S480	279–319	10–20	11	16, 17	15, 16	16, 17	20q13.2
D21S226	107–123	7–11	5	10, 10	10, 10	9, 10	21q21
D22S689	202–234	12–20	9	16, 16	16, 18	17, 17	22q12

[a] Product length includes seven bases sequence (tailed primer) and an adenine due to adenylation.

Results and discussion

We reviewed the literature and selected 10 polymorphic non-CODIS STR loci (D1S1656,[3] D2S1353,[4] D8S1132,[5] D12S1090,[6,7] D14S608,[8] D18S535,[9] D19S253,[10] D20S480,[11,12] D21S226,[12] and D22S689[12]) located on discrete chromosomes (Table 2). D1S1656, D14S608, D19S253, D20S480, and D22S689 are found on different chromosomes from the CODIS core loci. Although D2S1353 and D12S1090 are found on chromosomes forming part of the CODIS core loci (TPOX and vWA, respectively), the loci are located on different arms, and they are separated by a considerable distance. D8S1132, D18S535, and D21S226 are found on the same arms as the CODIS loci (D8S1179, D18S51, and D21S11, respectively), but these loci are more than 10 Mb apart from one another. Genetic

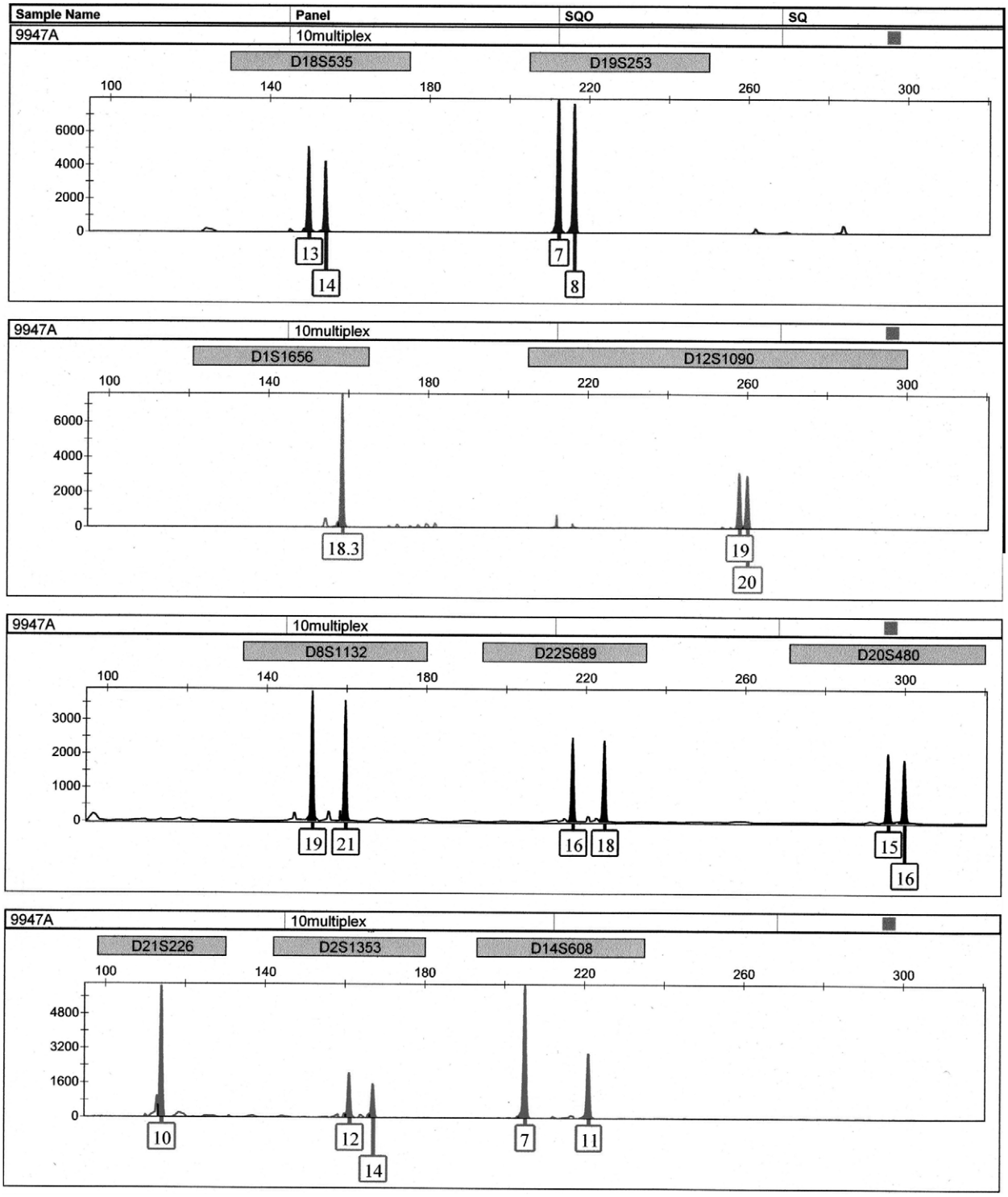

Fig. 1. Electropherogram of DNA typing result of control DNA 9947A using the present multiplex system for assay with 10 non-CODIS loci.

linkage disequilibrium (LD) in the human genome is generally reported at distances of 50–100 kb.[13–15] Thus, LD between the current non-CODIS eight loci and the CODIS core loci may be excluded from consideration. Analysis of extended STR loci using both the present multiplex system and commercial multiplex kits that include the CODIS core loci have been demonstrated to be remarkably effective in cases involving difficult kinship testing. The 10 non-CODIS loci examined in this study were analyzed using a newly developed multiplex PCR system that included 10 primer sets customized by Applied Biosystems. A 7-base sequence (tailed primer, Applied Biosystems) was appended to the 5′ end of the all reverse primers to promote adenylation.[16–18] PCR amplification and allele typing with the system was successful for all 10 loci. Fig. 1 shows an example of the typing of a commercially available cell line 9947A DNA. Very minor artifact peaks were detected, which had no effect on allele typing. Following analysis of the sensitivity range of the DNA detection limit, allele typing at all 10 loci proved successful when using at least 400 pg of template DNA with peak heights exceeded 100 RFU. In an analysis with 200 and 300 pg of template DNA, typing at all 10 loci was successful due to an increase in the amount of Taq polymerase (2.5 U). The amplification products for the 10 loci from 252 Japanese individuals ranged in length from 107 bp for D21S226 to 319 bp for D20S480. Numbers of alleles ranged from 5 alleles for D21S226 to 20 alleles for D12S1090. Table 2 provides information on the 10 loci and results for the genotype of a cell line DNA, while the whole genotyping data of 252 Japanese individuals are shown in Supplementary Table 1. Previous papers have used two styles of allelic nomenclature for D20S480. One paper refers to the repeat motif of allele (n) as $(TCTA)_n$,[11] while the other refers to $(TATC)_{n-2}ATC(TATC)_2$.[12] The present study followed the latter convention for nomenclature.

Table 3
Allele frequencies and statistical parameters of the 10 loci in a Japanese population ($N = 252$)

Allele	D1S1656	D2S1353	D8S1132	D12S1090	D14S608	D18S535	D19S253	D20S480	D21S226	D22S689
6					0.040		0.002			
7					0.192		0.153		0.008	
8		0.002		0.004	0.026		0.046		0.087	
9		0.002		0.008	0.139	0.234	0.010		0.143	
10		0.050		0.052	0.216	0.022	0.028	0.002	0.746	
11	0.028	0.018		0.133	0.206	0.008	0.121	0.002	0.016	
12	0.052	0.143		0.117	0.135	0.131	0.333	0.079		0.026
13	0.109	0.290		0.038	0.044	0.270	0.246	0.052		0.006
14	0.067	0.139		0.020	0.002	0.238	0.058	0.117		0.012
15	0.234	0.175				0.077	0.004	0.222		0.133
16	0.284	0.083	0.004			0.016		0.218		0.240
16.3	0.012									
17	0.087	0.095	0.034	0.004		0.002		0.228		0.363
17.3	0.063									
18	0.014	0.004	0.212	0.004		0.002		0.069		0.161
18.3	0.040									
19	0.004		0.196	0.008				0.008		0.052
19.3	0.006									
20			0.165	0.002				0.002		0.008
21			0.117	0.006						
22			0.143	0.014						
23			0.099	0.032						
24			0.028	0.069						
25			0.002	0.147						
26				0.157						
27				0.141						
28				0.038						
29				0.008						
OH[a]	0.825	0.821	0.845	0.865	0.877	0.798	0.766	0.766	0.413	0.778
PD[b]	0.952	0.946	0.952	0.976	0.946	0.922	0.916	0.945	0.623	0.908
PE[c]	0.647	0.639	0.686	0.725	0.749	0.595	0.537	0.537	0.122	0.558
PIC[d]	0.81	0.81	0.82	0.88	0.81	0.76	0.75	0.80	0.38	0.73
TPI[e]	2.86	2.80	3.23	3.71	4.06	2.47	2.14	2.14	0.85	2.25
P[f]	0.5799	0.1414	0.3366	0.5896	0.8610	0.4140	0.0660	0.1540	0.5509	0.8754

[a] Observed heterozygosity.
[b] Power of discrimination.
[c] Power of exclusion.
[d] Polymorphism information content.
[e] Typical paternity index.
[f] P-values of the exact tests for Hardy–Weinberg equilibrium.

Table 3 shows the observed allele frequencies and forensic statistical parameters for the 10 loci in the Japanese population. No significant deviations from Hardy–Weinberg equilibrium were observed in any of the 10 loci ($P > 0.05$). Heterozygosity values exceeded 0.76, except for D21S226 (0.413). The power of discrimination values ranged from 0.976 for D12S1090 to 0.908 for D22S689, except for D21S226 (0.623). All forensic parameters for D21S226 were lowest among the 10 loci, while the other 9 loci exhibited high polymorphism. The accumulated power of discrimination and power of exclusion for the 10 loci were 0.999999999998 and 0.99991, respectively. No mutations were detected in the kinship cases investigated (45 family trios). Given the infrequent meioses observed in the study, estimating a reliable mutation rate would require further investigation.

The previous literature report several earlier population studies of the present 10 non-CODIS loci from Asian populations. Single locus comparisons from this published data indicate that the distribution of D19S253 alleles is similar to those in Korean and Thai populations[19,20] (P values = 0.65 and 0.30, respectively). In D14S608, the Korean population[8] differed significantly from the present Japanese population ($P < 0.00001$). Comparisons of available Chinese published data[21,22] indicate that distributions of D2S1353 and D18S535 alleles also differ significantly from the Japanese data ($P < 0.00001$). On the other hand, published data on the German population exist for the eight loci,[6,12,23,24] except for D2S1353 and D19S253 and the distribution of alleles for D8S1132 alone showed no differences between German and Japanese populations ($P = 0.32$). However, significant differences between German and Japanese populations were found at the other loci ($P < 0.00001$).

We conclude that analysis of extended STR loci using the present multiplex PCR system for ten non-CODIS loci, when added to analysis of core CODIS loci, can be highly effective for difficult kinship testing, including deficient paternity cases.

Appendix A. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jflm.2008.04.001.

References

1. Maniatis T, Fritsch EF, Sambrook J. *Molecular cloning: a laboratory manual*. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory; 1989.
2. Raymond M, Rousset F. GENEPOP (version 3.4): populations genetics software for exact tests and ecumenicism. *J Hered* 1995;**86**:248–9.
3. Lareu MV, Barral S, Salas A, Pestoni C, Carracedo A. Sequence variation of a hypervariable short tandem repeat at the D1S1656 locus. *Int J Legal Med* 1998;**111**:244–7.
4. Leopoldino AM, Pena SD. The mutational spectrum of human autosomal tetranucleotide microsatellites. *Hum Mutat* 2002;**21**:71–9.
5. Wiegand P, Schneider HR, Schurenkamp M, Kleiber M, Brinkmann B. Tetranucleotide STR system D8S1132: sequencing data and population genetic comparisons. *Int J Legal Med* 1998;**111**:180–2.
6. Stein C, Lange T, Ferencik S, Grosse-Wilde H, Henβge C. German population data of three tetrameric short tandem repeat loci – D3S1744, D12S1090 and D18S849. *Forensic Sci Int* 1998;**91**:103–7.
7. Mertens G, Mommers N, Boutrand L, Gielis M, Vandenberghe A. Flemish population data and sequence structure of the hypervariable tetranucleotide repeat locus D12S1090. *Int J Legal Med* 2001;**115**:40–4.
8. Choi M, Kim JH, Lee DH, Lee SH, Rho HM. Frequency data on four tetrameric STR loci D18S1270, D14S608, D16S3253 and D21S1437 in a Korean population. *Int J Legal Med* 2000;**113**:179–80.
9. Lareu MV, Barral S, Salas A, Carracedo A. Sequence variation of a variable short tandem repeat at the D18S535 locus. *Int J Legal Med* 1998;**111**:337–9.
10. De Stefano F, Casarino L, Costa MG, Bruni G, Mannucci A, Unseld M, et al. Analysis of a short tandem repeat locus on chromosome 19 (D19S253). *Int J Legal Med* 1996;**108**:256–8.
11. Ohtaki H, Yamamoto T, Yoshimoto T, Uchihi R, Ooshima C, Katsumata Y, et al. A powerful, novel, multiplex typing system for six short tandem repeat loci and the allele frequency distributions in two Japanese regional populations. *Electrophoresis* 2002;**23**:3332–40.
12. Becker D, Vogelsang D, Brabetz W. Population data on the semen short tandem repeat loci D4S2366, D6S474, D14S608, D19S246, D20S480, D21S226 and D22S689 in a German population. *Int J Legal Med* 2007;**121**:78–81.
13. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, et al. Linkage disequilibrium in the human genome. *Nature* 2001;**411**:199–204.
14. Service SK, Ophoff RA, Freimer NB. The genome-wide distribution of background linkage disequilibrium in a population isolate. *Hum Mol Genet* 2001;**10**:545–51.
15. Weiss KM, Clark AG. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 2002;**18**:19–24.
16. Smith JR, Carpten JD, Brownstein MJ, Ghosh S, Magnuson VL, Gilbert DA, Trent JM, Collins FS. Approach to genotyping errors caused by nontemplated nucleotide addition by Taq DNA polymerase. *Genome Res* 1995;**5**:312–7.
17. Ginot F, Bordelais I, Nguyen S, Gyapay G. Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. *Nucleic Acids Res* 1996;**24**:540–1.
18. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* 1996;**20**:1004–10.
19. Shin CH, Lee JB, Lee YS. Allele frequencies and genetic data of 16 highly polymorphic loci in Koreans. *Forensic Sci Int* 2004;**141**:189–92.
20. Bhoopat T, Sriduangkaew S, Steger HF. STR loci D10S2325, D16S539 and D19S253: northern Thai population data. *Legal Med* 2006;**8**:306–7.
21. Sun XM, Fang YQ, Yun LB, Bai P, Li ZR, Zhang J. Population genetics of two STR loci D2S1346 and D2S1353 in a Han population of Chinese. *J Forensic Sci* 2006;**51**:705.
22. Tian YF, Zhao JH, Zhang HB, Zhu BF, Lai JH. Genetic profile of five STR loci D2S1338, D8S1179, D14S306, D19S253 and D18S535 in a Chinese population. *J Forensic Sci* 2005;**50**:245.
23. Wiegand P, Lareu MV, Schurenkamp M, Kleiber M, Brinkmann B. D18S535, D1S1656, and D10S2325: three efficient short tandem repeats for forensic genetics. *Int J Legal Med* 1999;**112**:360–3.
24. Schmid D, Anslinger K, Rolf B. Allele frequencies of the ACTBP2 (=SE33), D18S51, D8S1132, D12S391, D2S1360, D3S1744, D5S2500, D7S1517, D10S2325 and D21S2055 loci in a German population sample. *Forensic Sci Int* 2005;**151**:303–5.